THE UNIVERSITY OF QUEENSLAND
AUSTRALIA

Analysis of Engineering and Scientific Data

Semester 1 – 2019

Sabrina Streipert s.streipert@uq.edu.au

Geometric Distribution — Memoryless Property

Suppose, we have tossed the coin k times without a success (Heads).

What is the probability that we need more than x additional tosses before getting a success?

$$P(X = k{+}x \mid X = k) = \frac{P(X = x + k, X = k)}{P(X = k)} = \frac{P(X = x + k)}{P(X = k)} = \frac{p(1-p)^{k+x-1}}{p(1-p)^{k-1}} = (1{-}p)^x.$$

or similarly

$$P(X > k{+}x \mid X > k) = \frac{P(X > k + x, X > k}{P(X > k)} = \frac{P(X > k + x)}{P(X > k)} = \frac{(1-p)^{k+x}}{(1-p)^k} = (1{-}p)^x.$$

$\longrightarrow \underline{P(X > k + x \mid X > k) = P(X > x).}$ so the probability of having x more "losses" remains the same even if you had already k "losses"

Continuous Distribution I - Uniform

X has a **uniform distribution** on $[a, b]$, $X \sim \mathsf{U}[a, b]$, if its pdf f is given by

$$f(x) = \begin{cases} \frac{1}{b-a} & x \in [a, b] \\ 0 & \text{else} \end{cases}$$

For example used when modelling a randomly chosen point from the interval $[a, b]$, where each choice is equally likely.

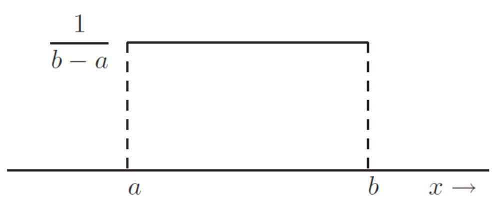


Figure 1: The pdf of the uniform distribution on $[a, b]$.

Uniform Distribution - Properties

- $\mathbb{E}[X] = \underline{\int_{-\infty}^{\infty} x f(x)\,\mathrm{d}x = \int_a^b x \frac{1}{b-a}\,\mathrm{d}x = \frac{1}{2(b-a)}(b^2 - a^2) = \frac{a+b}{2}}$

- $\mathrm{Var}(X) = \underline{E[X^2] - (E[X])^2 = \int_a^b x^2 \frac{1}{b-a}\,\mathrm{d}x - \left(\frac{a+b}{2}\right)^2 = \frac{\overbrace{(b^3 - a^3)}^{(b-a)(b^2+ab+a^2)}}{3(b-a)} - \frac{a^2+2ab+b^2}{4}}$

  $\underline{= \frac{b^2+ab+a^2}{3} - \frac{a^2+2ab+b^2}{4} = \frac{4b^2+4ab+4a^2-3a^2-6ab-3b^2}{12} = \frac{a^2-2ab+b^2}{12} = \frac{(a-b)^2}{12}}$

Example (Uniform Distribution):

Draw a random number from the interval of real numbers $[0, 2]$. Each number is equally possible. Let X represent the number.

1. Find the probability density function f.

$$f(x) = \begin{cases} \frac{1}{2} & x \in [0, 2] \\ 0 & \text{else} \end{cases}$$

2. What is the probability that the chosen number is within $(1.8, \pi)$?

   Note that $\pi > 2$!

$$\int_{1.8}^{2} \frac{1}{2} \,\mathrm{d}x = \frac{2 - 1.8}{2} = 0.1$$

Continuous Distribution II - Exponential

X is an **exponential distribution** with parameter $\lambda > 0$, $X \sim \mathsf{Exp}(\lambda)$, if the pdf of X is

$$f(x) = \begin{cases} \lambda \mathrm{e}^{-\lambda x}, & x \geq 0, \\ 0 & \text{else} \end{cases}$$

$\sim$ continuous version of the geometric distribution.

Used for example to analyse:

- Lifetime of an electrical component.

- Time between arrivals of calls at a telephone exchange.
- Time elapsed until a Geiger counter registers a radio-active particle.

Exponential Distribution - Properties

- Using integration by parts (Recall $\int_a^b v'w = vw|_a^b - \int_a^b wv$)

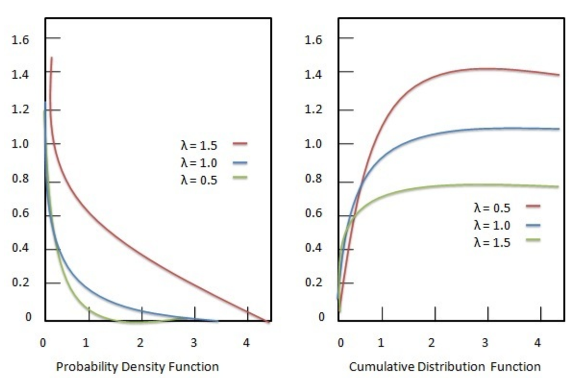
$$\mathbb{E}[X] = \int_0^\infty \overbrace{x}^{w} \overbrace{\lambda e^{-\lambda x}}^{v'} \,\mathrm{d}x = x\left(-e^{-\lambda x}\right)|_0^\infty - \int_0^\infty \left(-e^{-\lambda x}\right)\,\mathrm{d}x$$
$$= 0 - 0 - \int_0^\infty \left(-e^{-\lambda x}\right)\,\mathrm{d}x = -\frac{1}{\lambda}e^{-\lambda x}|_0^\infty = -0 + \frac{1}{\lambda} = \underline{\frac{1}{\lambda}}$$

- $\mathrm{Var}(X) = \underline{\frac{1}{\lambda^2}}$

- The cdf of X is given by:

$$F(x) = P(X \le x) = \int_0^x \lambda e^{-\lambda x}\,\mathrm{d}x = -\int_0^x \left(-\lambda e^{-\lambda x}\right)\,\mathrm{d}x = -e^{-\lambda x}|_0^x = 1 - e^{-\lambda x}$$

if $x \ge 0$, else $F(x) = 0$. (Note $F(0) = 0$ since the integral collapses.)



Exponential Distribution - Memoryless Properties

- The Exponential Distribution is the only continuous distribution that has the memoryless property:

$$
\begin{aligned}
P(X > s+t \mid X > s) &= \frac{P(X > s+t, X > s)}{P(X > s)} = \frac{P(X > s+t)}{P(X > s)} \\
&= \frac{1 - P(X \leq s+t, X > s)}{1 - P(X \leq s)} = \frac{1 - F(s+t)}{1 - F(s)} = \frac{1 - [1 - e^{-\lambda(s+t)}]}{1 - [1 - e^{-\lambda s}]} \\
&= \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}} = e^{-\lambda t} = 1 - [1 - e^{-\lambda t}] = 1 - F(t) \\
&= P(X > t)
\end{aligned}
$$

Example (Exponential Distribution):

The time a postal clerk spends with customers is exponentially distributed with the average time of 4 minutes. What is the probability that the clerk spends four to five minutes with a customer?

Since the mean is 4 minutes, we know that $\lambda = \frac{1}{4}$.

$$
P(4 \leq X \leq 5) = \underline{\int_4^5 \frac{1}{4} e^{-\frac{1}{4}x}\,\mathrm{d}x = e^{-\frac{1}{4}4} - e^{-\frac{1}{4}5} = e^{-1} - e^{-\frac{5}{4}} \approx 0.081}
$$

After how many minutes are half of the customers finished?

We need to solve $P(X \leq k) = 0.5$ for k:

$$
\begin{aligned}
P(X \leq k) &= 0.5 \\
F(k) &= 0.5 \\
1 - e^{-\frac{1}{4}k} &= 0.5 \\
0.5 &= e^{-\frac{1}{4}k} \\
\log(0.5) &= -\frac{1}{4}k \\
-4\log(0.5) &= k \qquad \longrightarrow k \approx 2.77
\end{aligned}
$$

Continuous Distribution III - Normal

X is **normal (or Gaussian) distributed** with parameters μ and σ^2, $X \sim \mathsf{N}(\mu, \sigma^2)$, if the pdf f of X is given by

$$
f(x) = \frac{1}{\sigma\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \quad x \in \mathbb{R}.
$$

Special case: Standard Normal Distribution

If $\mu = 0$ and $\sigma = 1$ then

$$
f(x) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}x^2}, \quad x \in \mathbb{R}
$$

Normal Distribution - Properties

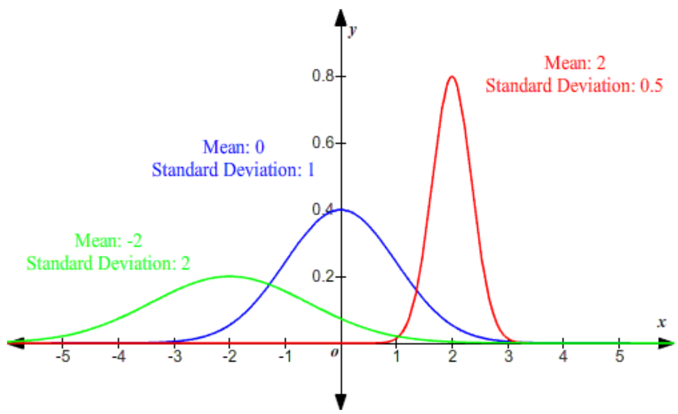
- $\mathbb{E}[X] = \underline{\mu}$
- $\mathrm{Var}(X) = \underline{\sigma^2}$

- If $X \sim \mathsf{N}(\mu, \sigma^2)$, then

$$\underline{\frac{X-\mu}{\sigma}} \sim \mathsf{N}(0,1). \qquad \longrightarrow \text{ standardisation}$$

- If $X \sim \mathsf{N}(\mu, \sigma^2)$, then

$$X = \underline{\mu + \sigma Z} \qquad \text{where } Z \sim N(0,1)$$



Source: https://www.varsitytutors.com/hotmath/hotmathhelp/topics/normal-distribution-of-data

It is very common to compute $P(a < X < b)$ for $X \sim \mathsf{N}(\mu, \sigma^2)$ via standardization as follows.

$$P(a < X < b) = P\left(\frac{a-\mu}{\sigma} < \overbrace{\frac{X-\mu}{\sigma}}^{Z \sim N(0,1)} < \frac{b-\mu}{\sigma}\right) = P\left(\frac{a-\mu}{\sigma} < Z < \frac{b-\mu}{\sigma}\right)$$

$$= P\left(Z < \frac{b-\mu}{\sigma}\right) - P\left(Z \leq \frac{a-\mu}{\sigma}\right) = F_Z\left(\frac{b-\mu}{\sigma}\right) - F_Z\left(\frac{a-\mu}{\sigma}\right)$$

That is,

$$P(a < X < b) = F_X(b) - F_X(a) = F_Z\left(\frac{b-\mu}{\sigma}\right) - F_Z\left(\frac{a-\mu}{\sigma}\right).$$

Inverse-Transformation

Exponential R.V. - inverse

For the exponential distribution, we have

$$F(x) = 1 - \mathrm{e}^{-\lambda x}.$$

For $y \in (0,1)$ and $\lambda \neq 0$,

$$y = 1 - e^{-\lambda x} \quad \leftrightarrow \quad \log(1-y) = -\lambda x \quad \leftrightarrow \quad -\frac{1}{\lambda}\log(1-y) = x$$

Chapter 5

1. Joint Probability Mass/Density Function
2. Joint Cumulative Distribution Function
3. Marginal Distributions, Marginal Expected Values
4. Covariance, Correlation Coefficient
5. Independence
6. Conditional pdf/pmf, Conditional cdf, Conditional Expectation
7. Generalization: Random Vector

Joint Distributions - Motivation

We analyse the student's weight X and height Y.

Aim: Specify a model for the outcomes?

Note: Can't just specify the pdf or pmf of one variable,, we also need to specify the interaction between weight and height

$\longrightarrow$ Need to specify the joint distribution of weight and height (in general: all random variables $X_1, \ldots, X_n$ involved).

Joint Distributions - Example

Example:

In a box are three dice.

- Die 1 is a normal die;
- Die 2 has no 6 face, but instead two 5 faces;
- Die 3 has no 5 face, but instead two 6 faces.

The experiment consists of selecting a die at random, followed by a toss with that die.

Let X be the die number that is selected, and let Y be the face value of that die.

$\longrightarrow$ joint probability mass function/table

Note:

$$P(X = x) = \sum_{y_j \in \Omega_y} P(X = x, Y = y_j)$$

x \ y	1	2	3	4	5	6	$\sum$
1	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{3}$
2	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{9}$	0	$\frac{1}{3}$
3	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$	0	$\frac{1}{9}$	$\frac{1}{3}$
$\sum$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	1

Figure 2: Joint pmf for unfair dice example.

and

$$P(Y = y) = \sum_{x_i \in \Omega_x} P(X = x_i, Y = y)$$

$\longrightarrow$ marginal probability (mass) function

1. $P(X = 2, Y = 4) = \frac{1}{18}$ (value in the 2nd row and 4th column in the probability mass table)
2. $P(X = 2) = \frac{1}{3}$ (sum all the probabilities in the second row)
3. $P(Y < 5) = \frac{4}{6} = \frac{2}{3}$ (sum all the probabilities in columns 1 to 4)

Joint Probability Mass Function

Definition:

Let (X, Y) be a discrete random vector. The function $(x, y) \to P(X = x, Y = y)$ is called the ***joint probability mass function*** of X and Y.

Let $\Omega_{X,Y}$ be the set of possible outcomes of (X, Y). Let $B \subset \Omega_{X,Y}$, then

$$P((X, Y) \in B) = \sum_{x \in B_X} \sum_{y \in B_Y} P(X = x, Y = y)$$

Joint pmf - Properties

1. $P(X = x, Y = y) \geq 0$, for all x and y.

2. $\sum_{x\in\Omega_X} \sum_{y\in\Omega_Y} P(X = x, Y = y) = 1$.

3. Marginal pmf of X:

$$P_X(X = x) = \sum_{y\in\Omega_y} P(X = x, Y = y)$$

4. Marginal pmf of Y:

$$P_Y(Y = y) = \sum_{x\in\Omega_x} P(X = x, Y = y)$$

Joint Probability Density Function

Definition:

We say that continuous random variables X and Y have a joint probability density function if, for all events $\{(X, Y) \in A \subset \Omega \subset \mathbb{R}^2\}$, we have

$$P((X, Y) \in A) = \int_{x\in A_x} \int_{y\in A_y} f_{X,Y}(x, y)\,\mathrm{d}y\,\mathrm{d}x$$

We often write $f_{X,Y}$ for f.

Joint pdf - Properties

1. $f(x, y) \geq 0$, for all x and y.

2. $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\,\mathrm{d}x\,\mathrm{d}y = 1$.

3. Marginal pdf of X:

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)\,\mathrm{d}y$$

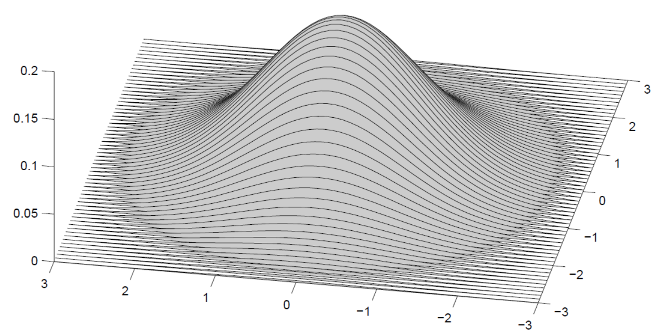
4. Marginal pdf of Y:

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\,\mathrm{d}x$$

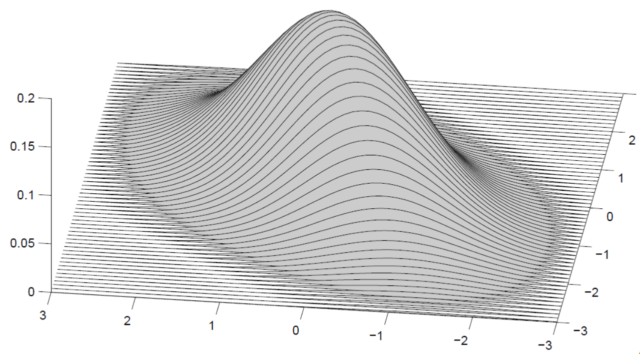
Example of a joint pdf: The Bivariate normal distribution:

We say that X and Y have a *bi-variate Gaussian (or normal) distribution* with parameters $\mu_1, \mu_2, \sigma_1^2, \sigma_2^2$ and $\rho \in [-1, 1]$ if the joint density function is given by

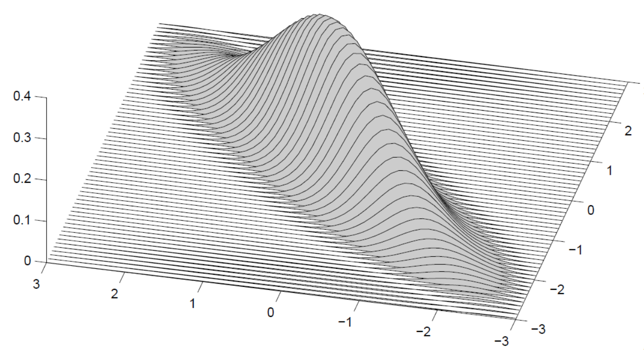
$$f(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{\frac{(x-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2}}$$



$\rho = 0.0$



$\rho = 0.5$



$\rho = 0.9$

Joint Cumulative Distribution Function

Joint Cumulative Distribution Function (joint cdf):

$$F(x, y) = P(X \leq x, Y \leq y)$$

- If X, Y are discrete R.V., then the joint cdf is

$$F(x, y) = \sum_{x_i \leq x} \sum_{y_j \leq y} P(X = x_i, Y = y_j)$$

- If X, Y are continuous R.V., then the joint cdf is

$$F(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(x,y) \, \mathrm{d}y \, \mathrm{d}x$$

Note also:

$$\frac{\partial^2}{\partial x \partial y} F(x,y) = f(x,y).$$

Expected value

- If X, Y are discrete R.V., **the marginal expected value** is:

$$E[X] = \sum_{x \in \Omega_x} x P_X(X = x) = \sum_{x \in \Omega_x} x \left[\sum_{y \in \Omega_y} P(X = x, Y = y) \right],$$

where $P_X(X = x)$ is the marginal pmf of X.

- If X, Y are continuous R.V., **the marginal expected value** is:

$$E[X] = \int_{x \in \Omega_x} x f_X(x) \, \mathrm{d}x = \int_{x \in \Omega_x} x \left[\int_{\mathbb{R}} f(x,y) \, \mathrm{d}y \right] \mathrm{d}x,$$

where $f_X(x)$ is the marginal pdf of X.

- Linearity:

$$E[aX + bY] = aE[X] + bE[Y]$$

Example - revisited

Recall the joint pmf for unfair dice example, see Figure 2.

1. Joint pmf: $P(X \leq 2, Y < 5) = \frac{8}{18}$ (sum all probabilities in columns 1 to 4 and rows 1 to 2.)

2. Marginal expected value:

$$E(X) = \sum_{i=1}^{3} iP_X(X = x) = \sum_{i=1}^{3} i\frac{1}{3} = \frac{1}{3}\frac{3 \cdot 4}{2} = 2$$

(since P_X is the marginal probabilities for X = summations over the rows)

3. Marginal expected value:

$$E(Y) = \sum_{j=1}^{j} P_Y(Y = y) = \sum_{j=1}^{6} j\frac{1}{6} = \frac{1}{6}\frac{6 \cdot 7}{2} = \frac{7}{2}$$

(since P_Y is the marginal probabilities for Y = summations over the columns)